Asian Research Journal of Information Technology, 2024
E-ISSN: 3028-0869
Volume 1 No. 2 P-ISSN: 3028-0850

Research on Blockchain-Based Industrial Internet Data Internet Data Security Protection

Huixin Luo, Wei Zhang, Bo Fan
https://orcid.org/0009-0009-6862-3101
https://orcid.org/0009-0001-4629-4257
https://orcid.org/0009-0002-6410-4821
University of the East
Industrial Information Security (Sichuan) Innovation Center Co., Ltd（icics）, Chengdu,610039, China
luo.huixin@ue.edu.ph, zhang.wei@ue.edu.ph, fan.bo@ue.edu.ph

ABSTRACT

With the development of industrial internet data informatization, privacy issues in the sharing and access of industrial internet data have attracted widespread attention from researchers. Blockchain, as decentralized, anonymous, and tamper-proof distributed ledger technology, provides a new approach to addressing privacy protection issues in industrial scenarios. This study first identified industrial internet data's privacy protection requirements and introduces the overall blockchain architecture. It then divides industrial information privacy protection into data-oriented privacy protection and user-oriented privacy protection and provides detailed descriptions of blockchain privacy protection methods for both data-oriented and user-oriented privacy protection. Data-oriented privacy protection focuses on safeguarding sensitive information itself, mainly using privacy protection methods based on encryption, distortion, and restricted release. User-oriented privacy protection focuses on protecting data users' privacy, including access control and transaction anonymity. Finally, we compare and summarize the characteristics of various methods, discuss the current research status of blockchain in privacy protection, and explore prospects for blockchain's development in the industrial information privacy protection domain.

Keywords: *Blockchain, Industrial information, Data privacy, User privacy, Decentralization*

INTRODUCTION

The rapid development of information technology is gradually transforming industrial services towards digitalization and informationization. We analyze and mine industrial Internet data as a crucial data asset, significantly advancing research and progress in the industrial field. Information exchange and sharing among various organizations can enhance the value of industrial Internet data, leading to a growing demand for industrial information sharing across various regions and organizations. Diverse sources and scopes of industrial Internet data hold special sensitivity and importance, encompassing not only health status and industrial processing information, but also personal privacy protection, industry development, and even national security. However, with the explosive growth of data and the application of deep

mining, the risk of leakage of industrial Internet data is increasing. We primarily categorize industrial Internet data leakage into two types: non-interactive leakage and interactive leakage. Non-interactive leakage pertains to the disclosure of internal systems or personnel, including the private use of permissions for information sale or misuse, while interactive leakage involves the release and sharing of industrial information among various organizations. Because industrial Internet data holds significant value in a centralized information system, the central node is motivated to protect privacy. During the process of multi-party interaction and sharing, all types of industrial Internet data are vulnerable to attacks, which can result in privacy leakage. Currently, the information technology department is responsible for the storage of industrial Internet data. However, many information systems lack security defense capabilities, technical measures are insufficient, and any attack on the system can result in the leakage of private information. Cloud computing platforms can provide powerful computing power for industrial Internet users, but the service model of directly sharing data raises a number of privacy and security issues, as well as data ownership issues. Industrial Internet data: Statistical analysis, data mining, deep learning, and other methods can filter out valuable information from massive data by eliminating meaningless parts. However, the processing of data, such as attribute matching and information correlation, could potentially expose sensitive information. Blockchain, a type of public decentralized distributed ledger, possesses features such as multi-party maintenance and non-tampering, making it ideal for addressing privacy concerns in data sharing on the Industrial Internet. It also facilitates the creation of data silos, offers a secure platform for sharing and trading data on the Industrial Internet, clarifies data ownership, and effectively guards against malicious tampering, abuse, and resale of data. The development and application of blockchain technology have brought new security and privacy issues, as well as convenience for industrial information sharing. Compared with centralized architecture, blockchain can avoid security risks caused by single points of failure or data leakage, thus ensuring data integrity and non-tampering. However, the transparency of transaction records in blockchain systems significantly increases the risk of privacy leakage, as analysis of these records can reveal user transaction patterns. In traditional industrial information systems, privacy protection focuses on ensuring the security of the central information system; however, in blockchain there is no unified administrator, and the messaging and consensus mechanisms employed also bring new opportunities and challenges for privacy protection.

METHODOLOGY

Background Knowledge of blockchain-based Privacy Protection

Industrial information privacy protection is required

With the continuous development of industrial informatization construction, the sharing and integration of industrial information provide users with better service quality. Numerous industrial Internet data sources store complex data types in heterogeneous information systems. Data sharing enables data interaction between different industrial information systems. As industrial services and research become more informalized, the issue of industrial information privacy also emerges. The high value and privacy of industrial information necessitates the

combination of privacy protection methods during information sharing, enabling the utilization of data value while maintaining clear data ownership and preventing the disclosure of sensitive information. To ensure the secure sharing of industrial Internet data, the data must satisfy the following requirements:

Privacy needs

Industrial Internet data sharing primarily concerns the privacy of data and related parties. Data sharing must limit the scope of sharing to prevent the leakage of private data to non-authorized parties, as illegal trading or use of private data will have immeasurable consequences.

Integrity requirements

Industrial Internet data users need to verify the data's reliability, including its source, tampering, forgery, and other factors. Incomplete or altered data will affect the efficiency of data sharing and may cause serious problems for the data user.

Availability Requirements

Availability refers to any authorized user's ability to access and use data at any time. Lack of availability in industrial internet data sharing means the provider will not be able to provide data services and the user will not be able to get the required data; therefore, ensuring data availability in the process of industrial sharing is the only way to ensure the benefits of data sharing and the value of industrial internet data.

Blockchain technology

Blockchain is a decentralized shared ledger represented by Bitcoin, a new technological architecture derived from digital cryptocurrencies. Cryptographic techniques ensure the unalterability and unforgeability of the chain's data, as blockchain composes blocks of data into a chained data structure in chronological order. The Hash algorithm and the Merkle tree enable each node to encapsulate transactions received over time into a time-stamped block, which it then links to the longest main chain to form the latest block. Every block consists of two components: the block header and the block body. The block header encompasses the previous block's hash value, a timestamp, a random number, and the Merkle root, among other elements. The block body contains the transaction data that the blockchain records. The Merkle tree hashing process generates a unique Merkle tree root, which the counterparty digitally signs and stores in the block header. The counterparty digitally signs each transaction and permanently stores it in the block for all users to query. While there are many commonalities in the overall architecture of blockchain systems, they are also constantly evolving. Generally, blockchain comprises a data layer, network layer, consensus layer, incentive layer, contract layer, and application layer; however, its development and evolution have weakened some traditional modules. Therefore, we can examine the fundamental characteristics of blockchain.

Network Layer

The network layer of the blockchain system is responsible for managing message propagation between nodes, data validation mechanisms, and ensuring all nodes participate in

block generation and verification. By utilizing a peer-to-peer (P2P) network structure, each node in the blockchain system can dynamically access and share information with other nodes. Each node broadcasts a newly generated transaction to the network, verifies its validity, and then forwards it to neighboring nodes in chronological order. This decentralized approach minimizes the risk of invalid transactions spreading throughout the network, enhancing security by preventing attackers from identifying the communication relationships between users through network traffic analysis.

Transaction Layer

The transaction layer facilitates the exchange of data between two addresses within a blockchain system, encompassing elements like address structure, transaction format, global ledger, and smart contracts. In the blockchain realm, addresses serve as user pseudonyms, typically generated via public key encryption algorithms. The public key functions as the input or output address for transactions, while the user safeguards the private key for transaction signing. Transactions log data interactions between addresses, with two predominant models: transaction-based (e.g., Bitcoin) and account-based (e.g., Ethereum, Hyperledger Fabric). The latter offers greater programmability and flexibility, offloading computational burdens from the chain to wallets, thereby alleviating pressure on the network. The choice between models should align with specific business needs. The global ledger archives all transaction data, contracts, and parameters in block form, with each block containing specific transaction details. Smart contracts, essentially programmable protocols deployed on blockchains, enable self-execution and verification triggered by external events. The global ledger records the details of smart contracts, just like it does transactions.

Consensus Layer

The consensus mechanism stands as the linchpin of blockchain systems, crucial for maintaining data security. A big problem is getting everyone to agree on how to keep the global ledger accurate. We employ various methods such as proof of work (POW), proof of stake (POS), delegated proof of stake (DPOS), and practical Byzantine fault tolerance (PBFT) to address this issue. POW, employed by Bitcoin, relies on computational competition to ensure data consistency. For node selection, POS replaces computational power with stake age, reducing resource waste. DPOS, a POS derivative, allows nodes to designate trusted proxies for block generation, enhancing transaction efficiency. PBFT, rooted in Byzantine fault tolerance, elevates fault tolerance rates and reduces algorithmic complexity. In spite of these mechanisms, the blockchain field is always changing, giving rise to new combined and variant mechanisms designed for different business situations, each with its own efficiency and security concerns.

Application Layer

The application layer encompasses a variety of scenarios, such as industrial Internet data sharing platforms, encapsulating programs, and user interaction interfaces. This layer, akin to decentralized applications (DApps), includes smart contracts and their invocation interfaces. Regulatory challenges loom over blockchain application development, balancing privacy

concerns with the need for oversight. Regulatory technology aims to detect and deter illicit activities, yet blockchain's decentralized, tamper-resistant, and pseudonymous nature complicates regulatory efforts, necessitating innovative approaches to ensure platform integrity.

Blockchain Privacy Protection

The core goal of data privacy protection is to protect data privacy and integrity. Preventing unauthorized users from accessing data is the definition of privacy, while ensuring the accuracy and lack of tampering is the definition of integrity. The blockchain's limited storage capacity necessitates the use of blockchain technology for data privacy protection, typically through separation of storage and management. The blockchain stores a significant amount of raw data on off-chain servers, managing data indexing and permissions, and ensuring data integrity through its open ledger. The blockchain data, visible to all users and impervious to tampering, enables access control technology to manage user rights. Unlike the authorization strategy that relies on a single node, the blockchain-based approach to managing user rights allows for open and transparent authorization and access control, while also preventing any potential overstepping by third parties. Blockchain's non-tampering characteristics enable the monitoring of data flow throughout the entire process, ensuring traceability and auditability in the data storage, transmission, and calculation processes. However, the openness and transparency of the global blockchain ledger will have some unfavorable effects on the privacy of data on the chain. Blockchain divides privacy into identity privacy and transaction privacy. [Identity privacy refers to the association between user identity information and blockchain address. The user's public key generates the address in the blockchain, which is independent of the user's identity information and can serve as the user's pseudonym within the system. However, the blockchain's anonymity does not ensure complete privacy, and when a user engages in a transaction using a blockchain address, they risk disclosing their identity information. Transaction privacy refers to the transaction data stored in the blockchain ledger and the knowledge behind the transaction data. The Bitcoin system records transactions publicly on the chain without taking any privacy measures. Furthermore, the industrial sector, through the use of data analytics and mining technologies, frequently reflects a variety of knowledge in its transaction records. The key components of blockchain privacy protection are identity privacy and transaction privacy. Once the mapping relationship between user identity and blockchain address is compromised, it can cause serious harm to user privacy. Unlike traditional centralized systems, it is impossible to limit the dissemination of leaked privacy information by deleting the data stored in the global ledger in blockchain. Even if a hard fork forms a new chain, the consequences of data leakage remain irreversible. Therefore, blockchain should emphasize the protection of privacy information. Industrial information privacy protection includes the protection of data in the process of publishing, storing, exchanging, and analyzing. This paper categorizes industrial information privacy protection methods into two categories: data-oriented and user-oriented. Data-oriented privacy protection pertains to safeguarding sensitive information within industrial Internet data and the

data itself, while user-oriented privacy protection pertains to authorizing user access to data and protecting user information during data use, transactions, and other related processes.

Data-Oriented Privacy Protection

Three categories primarily comprise industrial Internet data-oriented Internet data privacy protection: data encryption, data distortion, and data release restriction.

1. Privacy protection relies on data encryption. Blockchain employs a hash algorithm and digital signature-based transaction storage mechanism to guarantee the authenticity and non-repudiation of data on the chain. However, in intricate industrial business scenarios, the privacy of transaction data and smart contracts necessitates the integration of other cryptographic techniques like homomorphic encryption and secure multi-party computation. Moreover, the blockchain-based privacy protection system, which relies on data encryption, requires the protection of sensitive information itself. Additionally, in the blockchain-based industrial Internet data sharing platform, the generated blockchain address solely allows for the transfer and payment of its digital assets using the corresponding private key. Therefore, possessing the private key grants control over the address's digital assets, and the security of the wallet and key storage method significantly influences the privacy and security of the data within the blockchain system.

2. Homomorphic encryption in the era of industrial big data, data mining, learning, analysis, and other operations can bring more value to patients, hospitals, and related institutions, such as in-depth understanding of disease and the realization of personalized services. Data privacy protection is a crucial prerequisite for data sharing on the Industrial Internet, particularly during the computation and processing stages. Rivest et al. originally proposed homomorphic encryption in the 197Os, a class of encryption algorithms with unique properties. Instead of just basic encryption operations, homomorphic encryption can also perform multiple computation operations between ciphertexts. This means that computing followed by decryption is the same as decrypting followed by computation, which is very important for information security. In industrial scenarios, data owners who do not have computing resources can entrust a third party to compute the data and return the computation results, but there is a risk of privacy leakage from the untrustworthy third party. According to the literature, the improved homomorphic encryption algorithm can guarantee computational security in the MapReduce framework, so that the user can control the data autonomously while ensuring the security of data computation and sharing. The transparency of blockchain transactions and smart contracts makes it easy to leak data privacy. Homomorphic encryption can ensure the privacy verification of transaction data and the computational security of smart contracts, avoiding the leakage of privacy in the process of data verification or operation. Literature proposes a blockchain privacy protection method based on homomorphic encryption, in which the verification node operates and verifies the encrypted user transaction information, which improves blockchain technology's data security and privacy. The literature addresses the issue of insurance companies viewing patients' plaintext data during industrial

insurance claims, by integrating homomorphic encryption with blockchain smart contracts to address the privacy leakage issue.

3. Blockchain can save relevant operational records of industrial Internet data to the public ledger, allowing users to view them at any time and ensuring data integrity. Not only does the integration of blockchain and cloud technology address the storage and performance bottleneck of blockchain, but it also effectively prevents the falsification and tampering of industrial Internet data at will. However, cloud services and blockchain are facing the problem of how to ensure data privacy in practical scenarios. To some extent, homomorphic encryption can solve the data privacy problem in the cloud computing process; the data owner can entrust the encrypted data to an untrusted cloud service provider to process the data without disclosing the privacy information. In blockchain, homomorphic encryption not only provides privacy protection but also allows access to the encrypted data on the chain at any time for auditing or other operations. Using homomorphic encryption to store data on the blockchain will preserve the benefits of both public and private chains while keeping the advantages of Ethernet intact.

Secure Multi-party Computing Industrial Internet Data Internet data mining analysis can make the data produce value-added value, especially the statistical analysis of related data and reflecting the collective characteristics of the summary information, which has a higher value. However, due to the fragmentation of industrial Internet data storage, the private data of a single industrial node has the problems of insufficient data volume, insufficient data dimensions, etc. The privacy and high value of industrial Internet data often lead to reluctance among data owners to share the original data, thereby reducing the data's value-added value. Therefore, safeguarding the data privacy of all relevant parties during the multi-party joint computation process is crucial for the sharing of industrial Internet data. The generalized cryptographic primitive, known as f(x,x...x)=(yy...,y)(1), enables multiple data owners to jointly compute the same function in a mutually untrusted distributed environment without disclosing their input data x. After the computation, each participant cannot know any other information except the computation result and their own input values. SMPC completes the collaborative computation task under the premise of ensuring the privacy of the participants' inputs, and its framework includes cryptographic techniques such as obfuscated circuits, secret sharing, homomorphic encryption, and so on." The main features of the SMPC technique are as follows:

1) Privacy. No participant has access to another participant's input data.

2) Correctness. Each participant receives the correct calculation result.

3) Terminability. Guaranteed output for a limited time;

4) Fidelity. Most of the participants perform calculations as specified.

Both blockchain and SMPC involve participants interacting based on specific rules (protocols); blockchain prioritizes the verifiability of the calculation results and guards against tampering, while SMPC prioritizes the privacy of the input data during the calculation process. Its goal is to obtain accurate calculation results while maintaining input confidentiality. You can use blockchain as a depository for SMPC and apply it to blockchain smart contracts for

joint computation on multi-party data, all while ensuring the privacy of all parties. Through an incentive mechanism, the blockchain can publicize the participants' credibility and encourage honest participants, thereby addressing the SMPC's honesty problem. BFR-MPC is a blockchain-based multi-party computation scheme. The blockchain public ledger maintains a public credibility system for all parties and encourages all parties to execute the established computation protocols through the incentive mechanism, which better ensures fairness and robustness. The literature constructs a punishment mechanism based on blockchain smart contracts and proposes a secure SMPC protocol. This protocol determines malicious participants' early termination behavior through a timeout mechanism and imposes economic punishment on them. Literature constructed a smart contract framework based on SMPC, which ensures input privacy and computational correctness in smart contract execution. The introduction of cloud computing introduces a diverse and complex external environment for SMPC execution; we can regard the cloud server as a third party not fully trusted. Participants upload their respective inputs to the cloud server using homomorphic encryption, which computes the homomorphic ciphertexts and returns the results, thereby guaranteeing data confidentiality. The SMPC framework utilizes homomorphic encryption as a fundamental protocol, combining it with SMPC to safeguard the confidentiality of each participant's input data. SMPC offers a privacy computation scheme for industrial Internet data, guaranteeing the confidentiality of input data from multiple parties. Meanwhile, blockchain technology records the industrial Internet data circulation log, ensuring data confidentiality. Blockchain technology is designed to record the logs of industrial Internet data circulation, provide verifiability and traceability for privacy calculation, and incentivize participants to offer real data and honestly fulfill the established calculation protocols. When you combine blockchain and SMPC, you can safely share value-added industrial Internet data and make sure that data transactions can be checked and tracked. This makes it easier to regulate each participant and find the right people to be responsible.

Key Management

Blockchain combined with cryptography can enable secure industrial Internet data sharing. Because modern cryptographic systems require encryption algorithms to be public, the security of cryptographic systems does not depend on the algorithm's confidentiality, but on the key's confidentiality. Therefore, for encrypted industrial Internet data, reasonable preservation of the private key is the key to ensuring data privacy. In the blockchain system, the private key serves as an asset, and the wallet acts as a carrier for storing the private key. Mastering the private key requires having a blockchain address behind the data assets, so the security of the private key is especially critical to data privacy. We can categorize the storage of private keys into cold wallets and hot wallets based on network accessibility, and custodial wallets and non-custodial wallets based on third-party representation. Cold wallets are more secure than hot wallets due to their inaccessibility over the Internet, thereby reducing the risk of hackers stealing private keys. Conversely, hot wallets, when connected to the Internet, allow for instant trading, making them more convenient and speedier for frequent traders. Escrow wallets are wallets where a third party controls the cryptocurrency on behalf of the user, without

providing the user's private key or storing it on a local server. If a user loses his private key, he can recover his account on the server side, thus favoring those who have easy access to the wallet. However, if the service provider freezes the user's assets for reasons such as system maintenance or know your customer (KYC), there is a risk that the hosted type of wallet will be subject to theft by hackers. Moreover, the exchange controls the user's assets, and it has the authority to freeze them for purposes like real-name control or anti-money laundering. In noncustodial wallets, the user fully controls and possesses their cryptocurrency, while the device stores the encrypted private keys. In non-custodial-type wallets, if the user loses his/her auxiliary word, he/she will not be able to recover the assets. The literature proposes a distributed public key scheme based on blockchain technology, in which the nodes in the blockchain network share the responsibility of key storage. Compared to the traditional public key system, the scheme that divides the realistic storage system into distributed blockchain nodes exhibits superior response performance and anti-interference ability. To address the challenges of distributed network key management and communication overhead, scholars have proposed a key management scheme based on blockchain, known as KMSBoB. This scheme designs the key management and transmission process within a distributed group network, with its core being a blockchain-based key management strategy in a distributed network environment. KMSBoB is a blockchain-based key management strategy in a distributed network environment. We combine the MT/CO protocol with blockchain multi-node mining to form a multi-node session key generation protocol in a distributed environment. KMSBoB transforms the traditional key management model into a blockchain-based operation mode, unifying the three processes of verifying the legitimacy of the key component, transmitting the key component, and sharing the key component within the entire security management node, also known as the Secure Management Node (SMN). The cure management node (SMN) mining process involves all SMN members verifying the legitimacy of block data and implementing dynamic security management of key data. This process reduces the communication overhead and security risk associated with key leakage between traditional autonomous domains.

Privacy Protection Based on Data Restriction Release

In blockchain technology, there is an "impossible triangle." The so-called "impossible triangle" refers to the blockchain public chain; it is difficult to achieve decentralization, security, and high transaction processing performance at the same time. The blockchain's on-chain and off-chain systems are subject to certain limitations. While the off-chain system offers scalable computing and storage performance, it poses challenges in data verification to prevent tampering. On-chain blockchain ensures information transparency and tamper-proofness, but scalability and throughput are issues. Industrial Internet data usually has a large amount of information, complex types, and different needs for privacy protection, so it is not suitable to store a large amount of raw data or data with high privacy requirements directly in the public ledger. In real-world industrial settings, we can integrate the off-chain information system with blockchain; this ensures the integrity of the original industrial Internet data within the chain while also enabling the off-chain system to handle massive data storage and calculations.

Additionally, by limiting the release of on-chain information, the off-chain system enables the direct removal of sensitive industrial information from the public ledger, thereby safeguarding the security of private data. The off-chain storage stores the original data in the off-chain database, while the distributed ledger stores the corresponding digest hash value or index information. The hash algorithm can map any length of the string into a fixed-length binary value string using the original data mapping after the binary value string, which is the hash value. The large amount of raw data in the data-sharing program, which may contain private information, makes it unsuitable for chain storage. Therefore, the use of on-chain index table information and off-chain database storage not only frees up significant storage space on the blockchain but also enhances the efficiency of information sharing, thereby ensuring data security. large-scale industrial data does not lend itself to transmission during the consensus process. Consequently, the majority of current industrial blockchain systems employ this type of storage, exemplified by the literature's proposed blockchain-based industrial Internet data-sharing model (medical data share mode). While off-chain storage offers robust privacy, it's crucial for all parties to uphold off-chain database security, as the original data remains offline. Furthermore, the rise of quantum computing poses a threat to the security of hash algorithms, necessitating the long-term development of cryptographic technology for blockchain that can withstand quantum attacks.

User-Oriented Privacy Protection

User-oriented privacy protection refers to the user's access control rights to the data and its anonymity in the blockchain system. One of the most crucial methods for safeguarding data privacy is the implementation of access control policies, which serve to restrict the extent of data sharing and prevent unauthorized access to the information. Even though a user's identity in the blockchain system is different from their real-life identity, the processing of transaction records on the chain through analysis, clustering, and other methods makes it possible to map transaction relationships between different accounts and guess about the relationship between the input and output of transactions, which puts traders' privacy at risk. The blockchain's openness and transparency make anonymity a primary concern when storing industrial privacy information on the network. User-oriented privacy protection methods are mainly based on access control and transaction anonymization. Smart contracts and attribute encryption are two types of access control. There are many ways to hide transactions, such as hybrid protocols, zero-knowledge proofs, digital signatures, secure channel protocols, and K-anonymization techniques.

Privacy protection based on access control is critical in today's digital society. Smart contracts are a type of code that automatically executes contract conditions, allowing for secure handling of data and transactions. A smart contract effectively protects a user's personal information by not directly storing the user's sensitive data, instead relying on the code logic to ensure privacy. In addition, attribute-based encryption (ABE) technology further enhances privacy protection. Through attribute-based encryption, users can access specific encrypted data based on their specific attributes or permissions, thus realizing personalized access control.

This attribute-based encryption makes the risk of leaking private information much lower, and users can share data and participate in digital transactions more securely.

Smart contracts are automated contracts based on blockchain technology that can execute, manage, and verify contract terms without third-party intervention. Smart contracts use programming code to define contract conditions and rely on network nodes for validation and execution, enabling automated contract processing. Key features of smart contracts include automated execution, decentralization, and high security, making them applicable in various fields such as finance, real estate, and supply chain management. By encrypting data attributes, attribute encryption protects data during transmission, storage, and processing. Attribute encryption technology effectively secures personal privacy information, business secrets, and sensitive data, preventing unauthorized access and disclosure. Identity verification, data sharing, secure communication, and other areas widely utilize attribute encryption technology, which provides essential safeguards for information security. Privacy protection, based on transaction anonymity, encompasses a range of technical measures designed to safeguard participants' private information in digital currency transactions, thereby ensuring both privacy and security. This includes coin mixing mechanisms, zero-knowledge proof, digital signatures, secure channel protocols, and K-anonymity technologies.

First, the coin-mixing mechanism is a way to combine multiple transactions, making it extremely difficult to discern the relationship between transactions. Through the coin-mixing mechanism, transaction participants can conceal their real identities and transaction records, thus protecting personal privacy.

Second, zero-knowledge proof is a way to prove the validity of a statement without revealing its specific content. In digital currency transactions, zero-knowledge proof ensures privacy protection by proving the validity of a transaction without revealing related privacy information.

In addition, digital signatures, as an encryption technology, play a crucial role in digital currency transactions. Digital signatures can ensure transactions' authenticity and integrity, prevent tampering, and protect transaction participants' privacy.

Secure channel protocol refers to the construction of a secure channel during digital currency transactions.

RESULTS AND DISCUSSION

Future Research Challenges and Directions

Blockchain technology has yielded numerous research findings and offers distinct benefits in safeguarding privacy. However, it still faces numerous challenges in terms of security, privacy, transaction performance, and integrating with other privacy protection technologies. These are critical issues that must be addressed when implementing blockchain in the industrial information privacy protection domain.

1) On-chain data release.

Individuals or institutions cannot manipulate the data in the system due to blockchain's decentralization and tamper-proof characteristics, and even if a single node fails or experiences an attack, it won't affect the integrity of the data or the system's overall operation. Industrial information is characterized by its wide source, large data volume, complex data type, and strong privacy, among other things. The release and sharing of data can leak the privacy information contained in the industrial Internet data set. If this data is directly released on the chain without combining it with other privacy measures, it will inevitably impact data privacy. The limited block capacity, transaction performance, and throughput, which are major bottlenecks of blockchain technology, make it challenging for the blockchain to directly store large-scale data. The combination of on-chain and off-chain is inevitable for the large-scale storage and computational processing of raw industrial Internet data, necessitating the release of representative data from the blockchain ledger. By ensuring non-disclosure of privacy, this ledger guarantees the completeness of the data stored under the chain and the verifiability of the data transactions in real-world application scenarios. In addition, realizing the secure docking between the blockchain and the traditional information system under the chain is a key issue for further research.

2) Balance between anonymity and regulation.

Blockchain supports anonymous transactions, but from a practical perspective, it lacks true anonymity. As data analysis technology advances, attackers can still extract information from public transactions on the chain, using data mining and other technologies, and infer the correlation between the addresses. Data mining and other techniques enable the attacker to deduce the correlation relationship between addresses from the public transaction information on the chain, thereby compromising the guarantee of user privacy. From a negative perspective, this will significantly undermine the anonymity of the transaction. The transparency of the transaction will expose the user's identity, leading to the leakage of chain information. However, from a positive perspective, this will assist the regulatory mechanism in identifying traces of illegal transactions and crimes, as well as locating and tracking the responsibilities of relevant organizations and responsible individuals. In practical application, we should balance the relationship between anonymity and reliability with specific scenarios to protect the privacy of legitimate users and prevent illegal behaviors.

CONCLUSION

5 Research Directions of Blockchain in the Field of Industrial Information Privacy Protection.

1) Combination of off-chain privacy calculation and on-chain depository.

Raw industrial data serves as the foundation of the industrial information industry, but its value is significantly lower than the value-added value derived from relevant computation

and analysis processes (such as big data mining and deep learning). Individual organizations typically encounter issues such as a lack of sufficient samples and a lack of richness in data dimensions, necessitating the addition of data from multiple sources during the calculation process. However, given the high value and privacy attributes of this data, all owners of industrial Internet data must safeguard the privacy of their data during joint calculations. If the industrial Internet data is provided directly to the demand side or for interaction between multiple parties, it is very easy to cause privacy leakage, so many data users are reluctant to provide data, resulting in a large amount of data not being able to make greater use of value. Separating the ownership and usage of data, achieving the joint privacy calculation of multi-party industrial Internet data locally without leaving the warehouse, and sharing the value-added data generated by the joint calculation with the demand side can reduce the risk of privacy leakage. Homomorphic encryption, SMPC, and other methods for realizing data privacy calculations of industrial Internet data are effective; however, these methods focus on data protection during the calculation process and cannot guarantee the authenticity of the data provided by the participants or the verifiability of the calculation results. Blockchain, as a distributed ledger technology jointly maintained by multiple parties and capable of effectively preventing repudiation, focuses on emphasizing the verifiability of calculations as well as preventing tampering with calculation results. Blockchain incentives also help to promote the fulfillment of the established computation protocols by each computation participant and increase the cost of malpractice. Therefore, in the future, we can endeavor to merge the current privacy calculation technology with blockchain technology. This will guarantee data privacy and controllability via off-chain privacy calculation. Additionally, the blockchain will store the necessary credentials during the calculation process, ensuring the verifiability of the calculation outcomes.

2) Realize smart contracts-based search authorization and access control of data.

Owing to the intricate nature of the vast industrial Internet data types and varying privacy requirements that render on-chain storage unsuitable, the majority of current industrial blockchain applications encrypt and store the original data in an off-chain database. The chain only retains a summary of the data, such as a hash value, to safeguard its integrity. To ensure the privacy and security of industrial Internet data in the sharing process, it is necessary to implement a secure search of off-chain data and set up reasonable access control permissions. The smart contract has the feature of automatic execution, which can authorize the user's access and record the authorization for the user and the data search log in the public ledger. The access control strategy based on smart contracts is conducive to avoiding power centralization, making the data authorization process transparent and open, and realizing more flexible and automated access control.

3) Realize the regulatory mechanism of "chain against chain" and the traceability of illegal transactions.

Privacy protection's sustainability depends on regulatory technology. Researching the privacy protection of industrial information necessitates the regulation and tracing of illegal behaviors. The anonymity and decentralization of the blockchain present higher regulatory requirements, particularly for the public chain. We can adopt the alliance chain + private chain

regulatory mechanism in the industrial blockchain system. Each organization maintains the private chain internally to store data-related information and introduces a regulatory node in the alliance chain to store data transaction information. This allows us to fully utilize the advantages of the private chain, such as its fast transaction speed and privacy, while also ensuring the order of data market transactions and preventing illegal transaction behaviors through the alliance chain. Additionally, the regulator must promptly identify illegal transactions while safeguarding the privacy of legitimate users. They must also trace and monitor the entire process of these illegal transactions, as well as the individuals responsible, to provide criminal evidence for any malicious behavior.

SUMMARY

Industrial Informa ionization not only streamlines research and services in this field but also introduces numerous new privacy and security concerns. Blockchain's decentralization, tampering, anonymity, traceability, and other characteristics provide a solution to guarantee the security and privacy of industrial information in the sharing process. In this paper, which focuses on the industrial information field, we systematically sort out the definition, architecture, and data privacy protection needs of blockchain. We then introduce data and user privacy protection methods based on blockchain technology, analyze and compare these methods, summarize the problems and challenges of the existing methods, and finally look forward to the future research direction in light of the shortcomings and challenges of the current state of research.

ACKNOWLEDGMENTS

The researcher would like to express their gratitude to key R&D projects in Sichuan Province+. The researcher would like to express their gratitude to the key R&D projects in Sichuan Province, specifically the Airport Luggage High-Speed Sorting Security Control System Based on Zero Trust Security +2022YFSY0005/Chengdu Science and Technology Innovation Project and the Industrial Internet Data Security Sharing and Privacy Protection Technology +2021-YF08-00151-GX, for their support in conducting this study.

REFERENCES

Das, D. (2018). Secure cloud computing algorithm using homomorphic encryption and multi-party computation. In *2018 International Conference on Information Networking (ICOIN)* (pp. 391-396). IEEE.

Dong, X. Q., Guo, B., Shen, Y., & Others. (2018). An efficient and secure decentralizing data sharing model. *Chinese Journal of Computers, 41*(5), 1021-1036.

Feng, Q., He, D. B., Zeadally, S., et al. (2019). A survey on privacy preemption in blockchain system. *Journal of Network and Computer Applications, 126*, 45-58.

Gao, H. M., Ma, Z. F., Luo, S. S., et al. (2019). BFR-MPC: A blockchain-based fair and robust multi-party computation scheme. *IEEE Access, 7*, 110439-110450.

Han, X., Yuan, Y., & Wang, F. Y. (2019). Security problems on blockchain: The state of the art and future trends. *Acta Automatica Sinica, 45*(1), 206-225.

Hao, Q. F., Jin, C. Q., Zhang, Z., et al. (2018). Blockchain: Architecture and research progress. *Chinese Journal of Computers, 41*(5), 969-988.

Huang, J. H., Jiang, Y. H., & Li, Z. C. (2020). Constructing fair secure multi-party computation based on blockchain. *Application Research of Computers, 37*(1), 225-230, 244.

Jiang, H., & Xu, L. (2016). Secure multiparty computation in computing. *Journal of Computer Research and Development, 53*(10), 2152-2162.

Liu, A. D., Du, X. H., Wang, N., et al. (2018). Research progress of blockchain technology and its application. *Journal of Software, 29*(7), 2092-2115.

Nakamoto, S. (2020, August 10). Bitcoin: A peer-to-peer electronic cash system. *Bitcoin.org*. https://bitcoin.org/bitcoin.pdf (Original work published 2009)

Ouyang, J., Yin, J., & Liu, S. P. (2014). An elective differential privacy transaction data publication strategy. *Journal of Computer Research and Development, 51*(10), 2195-2205.

Ouyang, J., Yin, J., & Lu, S. P. (2015). Differential privacy publishing strategy for distributed transaction data. *Journal of Software, 26*(6), 1457-1472.

Wang, T., Ma, W. P., & Luo, W. (2019). Information sharing and secure multi-party computing model based on blockchain. *Computer Science, 46*(9), 162-168.

Xiong, P., Zhu, T. Q., & Wang, X. F. (2014). A survey on differential privacy and applications. *Chinese Journal of Computers, 37*(1), 101-122.

Xu, T. F., Fu, C. O., Wang, C., et al. (2017). A medical data sharing model via blockchain. *Automation, 43*(9), 1555-1562.

Yan, C. (2018). Survey on security of blockchain. *Journal of Gerontological Research, 5*(5), 458-469.

Yang, M., Margheri, A., Hu, R. S., et al. (2018). Differentially private data sharing in a cloud federation with blockchain. *IEEE Cloud Computing, 5*(6), 69-79.

Yang, Y. H., Wei, L. J., Wu, J., et al. (2020). Block-SMPC: A blockchain-based secure multi-party computation for privacy-protected shareholding. In *Proceedings of the 2020 The 2nd International Conference on Blockchain Technology* (pp. 46-51). ACM.

Yuan, Y., Ni, X. C., Zeng, S., et al. (2018). Blockchain consensus algorithms: The state of the art and future trends. *Acta Automatica Sinica, 44*(11), 2011-2022.

Zhao, C., Zhao, S. N., Zhao, M. H., et al. (2019). Secure multi-party computation: Theory, practice, and applications. *Information Sciences, 476*, 357-372.

Zhang, X. J., & Meng, X. F. (2014). Differential privacy in data publication and analysis. *Chinese Journal of Computers*, *37*(4), 927-949.

Zhu, Y., Song, X. X., Xue, X. B., et al. (2019). Smart contract execution system over blockchain based on secure multiparty computation. *Journal of Cryptologic Research*, *6*(2), 246-257.